SECTION II.—GENERAL METEOROLOGY.

PARTIAL CORRELATION APPLIED TO DAKOTA DATA ON WEATHER AND WHEAT YIELD.

By Thomas Arthur Blair, Observer. [Weather Bureau, Salt Lake City, Utah, Jan. 29, 1918.]

In papers previously published the writer has shown? (1) A definite relation between the rainfall of May and June and the yield of spring wheat in the Dakotas. This relation was expressed by the coefficients of correlation, $r = +0.63 \pm 0.05$ for North Dakota, and $+0.59 \pm 0.06$ for South Dakota.

(2) A negative relation of somewhat greater value between the June temperatures and the wheat yield, the coefficients being -0.67 ± 0.08 and -0.73 ± 0.07 , re-

spectively.

Since we have here three related variables capable of statistical statement, it seems advisable to extend the statistical statement, it seems advisable to extend the study of their mutual relations by the application of the method of "partial correlation." At the same time opportunity is taken to bring the tables down to date, and to present a brief explanation of the meaning and application of coefficients of "net" or "partial correlation."

The manner of constructing a correlation table and of computing coefficients of gross correlation has been frequently explained in recent issues of the Monthly Weather Review, is illustrated in Tables 1 and 2, herewith, and will need no further explanation. In addition to the correlations of yield with precipitation and temperature, there has been added in these tables a computation of the coefficient of correlation between precipitation and temperature. This shows for each State a rather high negative value. In other words, that the wet Mays and Junes in the Dakotas are in general the cool Junes. In view of this fact, the question naturally arises as to how much of the apparent relation between precipitation and yield, shown by the coefficients 0.611 and 0.487, is really due to the influence of precipitation, and how much is due to the simultaneous influence of temperature; and similarly, how much of the apparent relation between temperature and yield is due

Partial correlation coefficients.—Partial correlation coefficients enable us to answer such questions as these. In other words, they express the correlation between two related variables after eliminating the influence of one or more other variables, or, on the supposition that the other variables become constants. In the case of three quantities, the partial coefficient is given by the equation

$$r_{12\cdot 8} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^{2})(1 - r_{23}^{2})}}$$

In this equation, the variables are numbered 1, 2, and 3 and the terms r_{12} , r_{13} , and r_{23} are the ordinary coefficients of total correlation between the two variables indicated by the subscripts, and $r_{12'3}$ is the coefficient of partial correlation between the quantities 1 and 2 after eliminating the influence of 3. In the general case, for any number of variables, the equation becomes

$$r_{12\cdot 345\cdot \cdot \cdot \cdot \cdot n} = \frac{r_{12\cdot 45\cdot \cdot \cdot \cdot \cdot n} - r_{13\cdot 45\cdot \cdot \cdot \cdot \cdot n} r_{23\cdot 45\cdot \cdot \cdot \cdot \cdot n}}{\sqrt{(1 - r_{13\cdot 45\cdot \cdot \cdot \cdot \cdot n}^2)(1 - r_{23\cdot 45\cdot \cdot \cdot \cdot \cdot n}^2)}}$$

in which each of the terms represented by r is the partial correlation coefficient between the two variables before the period when the effect of the quantities following the period is taken into account. The numerical solution requires the successive application of these two equations eliminating one variable at a time.

In the case under consideration, we have only the three quantities, precipitation, temperature, and yield to consider, and representing these by the letters, p, t, and y, we can at once write the equations as follows:

$$\begin{split} r_{py \cdot t} &= \frac{r_{py} - r_{pt} \; r_{ty}}{\sqrt{(1 - r_{pt}^2)(1 - r_{ty}^2)}} \\ r_{ty \cdot p} &= \frac{r_{ty} - r_{pt} \; r_{py}}{\sqrt{(1 - r_{pt}^2)(1 - r_{py}^2)}} \end{split}$$

$$r_{pt\cdot y} = \frac{r_{pt} - r_{py} r_{ty}}{\sqrt{(1 - r_{py}^2)(1 - r_{ty}^2)}}$$

Substituting the values given in Tables 1 and 2, and solving, we have

For North Dakota:

$$r_{py\cdot t} = \frac{0.611 - (-0.376)(-0.448)}{\sqrt{\{1 - (-0.376)^2\}\{1 - (-0.448)^2\}}} = \frac{0.4425}{0.8283} = +0.534,$$

$$r_{ty\cdot p} = \frac{-0.448 - (-0.376)(0.611)}{\sqrt{\{1 - (-0.376)^2\}\{1 - (0.611)^2\}}} = \frac{-0.2183}{0.7335} = -0.298,$$

$$r_{pt\cdot y} = \frac{-0.376 - (0.611)(-0.448)}{\sqrt{\{1 - (0.611)^2\}\{1 - (-0.448)^2\}}} = \frac{-0.1023}{0.7078} = -0.144,$$

For South Dakota:

$$r_{py \cdot t} = \frac{0.487 - (-0.555)(-0.622)}{\sqrt{(1 - (-0.555)^2)(1 - (-0.622)^2)}} = \frac{0.1418}{0.6513} = +0.218,$$

$$r_{ty \cdot p} = \frac{-0.622 - (-0.555)(0.487)}{\sqrt{(1 - (-0.555)^2)(1 - (0.487)^2)}} = \frac{-0.3517}{0.7266} = -0.484,$$

$$r_{pt \cdot y} = \frac{-0.555 - (0.487)(-0.622)}{\sqrt{(1 - (0.487)^2)(1 - (-0.622)^2)}} = \frac{-0.2521}{0.6839} = -0.369.$$

We may summarize the results thus, dropping the third figure,

North Dakota:

$$r_{py} = +0.61;$$
 $r_{ty} = -0.45;$ $r_{pt} = -0.38;$ $r_{py,t} = +0.53;$ $r_{ty,p} = -0.30;$ $r_{pt,y} = -0.14;$

South Dakota:

$$r_{py} = +0.49;$$
 $r_{ty} = -0.62;$ $r_{pt} = -0.56;$ $r_{py,t} = +0.22;$ $r_{ty\cdot p} = -0.48;$ $r_{pt\cdot y} = -0.37.$

The coefficients of total correlation, it will be noted, are here somewhat smaller than in the original calculation for the shorter period.

Interpretation of results.-The actual relation between the data for precipitation and yield is shown by the coefficients +0.61 and +0.49; but these are reduced to +0.53 and +0.22 by eliminating the influence of temperature, and these latter thus become the expression of the real causative relation between precipitation as such

¹ Monthly Weather Review, October, 1913, 41: 1515-1517. Monthly Weather Review, January, 1915, 43: 24-26.

² See especially *Marvin*, *C. F.*: Elementary notes on least squares, etc. Monthly Weather Review, October, 1916, 44: 551-569.

and yield. A considerable part of the apparent effect of precipitation upon yield is thus seen to be due to the influence of the accompanying temperatures upon the yield. Similarly, the real effect of temperature alone is shown by the coefficients -0.30 and -0.48 instead of -0.45 and -0.62. The interrelation of the three factors is thus clearly shown; and, further, the values +0.53 and -0.30 show that for North Dakota the net effect of precipitation is greater than that of temperature, while, on the other hand, in South Dakota the influence of temperature is the greater, as shown by the coefficients +0.22 and -0.48. We are aware, of course, that the precipitation of May and June and the temperature of June are not the only influences determining the final yield of wheat, and this is indicated in these results by the fact that the values found for the net coefficients do not approximate unity. If we could take account of all the factors, we should be able to find a coefficient of partial correlation between the yield and any one of them which would be approximately +1 or -1. The magnitude of the coefficients here found is sufficient evidence that the precipitation and temperature expressed in montidy sums and means are two of the important factors affecting yield, but not the only important factors.

The values of r_{pt} disclose a meteorologically interesting connection between precipitation and temperature. which need not be here discussed. It is necessary, however, to notice the relative values of the coefficients r_{pt} and $r_{pt,n}$, for these illustrate an important point in interpreting the results of partial correlations, showing the necessity for care in arriving at the true significance of such functions. We note that the relations between precipitation and temperature, -0.38 and -0.56, are reduced to -0.14 and -0.37 when we consider yield. Now, to follow strictly the line of argument used in interpreting the other partial coefficients would lead to the reductio ad absurdum that the summer yield of wheat affected the relation between the temperature and precipitation of the previous May and June. The point to be noted is that in properly interpreting coefficients of partial correlation we must be able to determine from other considerations which of the variables are causes and which are effects. In the problem under consideration there is evidently no difficulty on this score; we know that yield is an effect and the other variables are causes. The relation actually obtaining between precipitation and temperature is expressed by the coefficient r_{pt} , and the meaning of r_{pt} , is that if the yield had been the same, a different relation would have subsisted between precipitation and temperature, which is but another way of arriving at the general conclusion that there is a real relation between precipitation, temperature, and yield.

CONCLUSIONS.

(1) The precipitation of May and June and the temperature of June are important factors, but not the only important factors, affecting the yield of wheat in the Dakotas.

(2) A considerable part of the apparent effect of either precipitation or temperature upon yield is really due to the accompanying effect of the other.

(3) In North Dakota the influence of precipitation is greater than that of temperature, while the reverse is true in South Dakota.

(4) When the precipitation of May and June is above the average in the Dakotas the temperature of June is generally below the average, and inversely.

Table 1.—Correlation between the rainfall of May and June, the temperature of June, and the yield per acre of spring wheat in North Dakota.

Year.	May and June rainfall.			June tempera- ture.			Yield.					
	Amt.	De- part. p.	μ².	Mean,	De- part. t.	ſ².	Amt.	De- part. y.	y².	pt.	py.	ty.
1892 1893 1894	5, 0 5, 0	-0.5 1.1 1.1 +1.1	1.21 1.21	67.4	+4.8 +6.2	23. 04 38. 44	9.6 11.8	-2.5 -0.3	6, 25 0, 09	+ 1.05 - 5.28 - 6.82 - 3.19	$\begin{array}{r} + 2.75 \\ + 0.33 \end{array}$	- 0.21 -12.00 - 1.86 -25.81
1896 1897 1898 1899 1900	4. 5 5. 2 7. 3	-1.6 -0.9 +1.2	2, 56 0, 81 1, 44	61. 7 62. 6 62. 2	-0.9 0 -0.4	0, 81 0 0, 16	10.3 14.4 12.8	+2.3	3. 24 5. 29 0. 49	+ 1.44 0 - 0.48	- 0.78 + 2.88 - 2.07 + 0.84 +28.80	+ 1.62 0 - 0.28
1901 1902 1903 1904	7.2 4.5	+0.4 $+1.1$ -1.6 $+1.4$ $+1.5$	1, 21 2, 56 1, 96	58. 0 62. 4 61. 4	-1.0 -4.6 -0.2 -1.2 -2.9	21, 16 0, 04 1, 44	15.9 12.7 11.8	+3.8	14, 44 0, 36 0, 09	- 5.06 + 0.32 - 1.68	+ 4.18 - 0.96	- 0.12 + 0.36
1906 1907 1908 1909 1910	4.0 6.3 6.5	+0.4	4.41 0.01 0.10	$\begin{array}{c} 61.9 \\ 00.4 \\ 62.9 \end{array}$	-2.2 +0.3	0, 49 4, 84 0, 09	10.0 11.6 13.7	+1.6	4, 41 0, 25 2, 56	+ 1, 47 - 0, 44 + 0, 12	$\begin{array}{c} +\ 2.34 \\ +\ 4.41 \\\ 0.10 \\ +\ 0.64 \\ +21.78 \end{array}$	+1.10 +0.48
1911 1912 1913 1914 1915	6.6 4.0 8.4	-2.1 + 2.3	0.25 4.41 5.29	61, 8 65, 8 62, 2	+4.3 -0.8 +3.2 -0.4 5.9	0, 64 10, 24 0, 16	18.0 10.5 11.2	+5.9 -1.6 -0.9	34. 81 2. 56 0. 81	— 0, 40, — 6, 72 — 0, 92	+2.95 +3.36	-17.63 - 4.72 - 5.12 + 0.36 -35.99
1916 1917 Sums.	2.4	-3.7 -4.7	13, 69 89, 91	58.5	-4.4 -4.1	16, 81 263, 19	8.2	-3.9 -4.9	368. 57	+15, 17 55, 46	+14.43	+29, 04 +15, 99 - 140, 73
Means		0.18	ļ	62.6	—0. 17			-0.2ŏ			100.00	140.10

Table 1-computations.

$$[p] = -4.7; \quad [t] = -4.5; \quad [y] = -4.9; \quad [p^2] = 89.91; \quad [t^2] = 263.19; \quad [y^2] = 368.57; \quad [pu] = 109.50; \quad [ty] = -140.73; \quad [pt] = -55.46; \quad [pt] = -5$$

 $Er_{pt} = \pm \frac{0.6745}{5.1} \times 0.8586 = \pm 0.113;$

TABLE 2.—Correlation between the rainfall of May and June, the temperature of June, and the yield of spring wheat per acre, in South Dakota.

	•	•		•	•			•	-			
	Precipitation.			Temperature.			Yield.					
Year.	Amt.	De- part. p.	p^2 .	Mean.	De- part. t.	p.	Amt.	De- part. y.	y ².	pt.	py.	ty.
1891 1892 1893 1894	9.5 4.5 3.7	-0.3 +2.7 -2.3 -3.1 +0.1	0. 09 7. 29 5. 29 9. 61 0. 01	63.9 70.3 70.6	-1.7 -2.0 +4.4 +4.7 -2.2	2. 89 4. 00 19. 36 22. 09 4. 84	12. 5 8. 5 6. 6	+4.0 +1.3 -2.7 -4.6 +0.8	1. 69 7. 29 21. 16	+0.51 -5.40 -10.12 -14.57 - 0.22	+3.51 +6.21 +14.26	$ \begin{array}{r r} -2.60 \\ -11.88 \\ -21.62 \end{array} $
1896 1897 1898 1899 1900	4.6 6.8 8.1	-0.2 -2.2 0 +1.3 -3.3	0. 04 4. 84 0 1. 69 10. 89	65. 0 67. 3 66. 4	+1.1 -0.9 +1.4 +0.5 +3.5	1. 21 0. 81 1. 96 0. 25 12. 25	8.0 12.4 10.7	0 -3.2 +1.2 -0.5 -4.3	0 10. 24 1. 44 0. 25 18. 49	+1.98 0 +0.65	+7.04	+2.88 +1.68 -0.25
1901 1902 1903 1904 1905	6.0 7.0 6.5	+1.3 -0.8 +0.2 -0.3 +4.8	1. 69 0. 64 0. 04 0. 09 23. 04	62. 6 65. 0 64. 5	+0.4 -3.3 -0.9 -1.4 -1.5	0. 16 10. 89 0. 81 1. 96 2. 25	12. 2 13. 8 9. 6	+1.7 +1.0 +2.6 -1.6 +2.5	2. 89 1. 00 6. 76 2. 56 6. 25	+2.64 -0.18 +0.42	-0.80 +0.52	-3.30 -2.34 +2.24
1906 1907 1908 1909	7.7 10.0 9.0	+1.6 +0.9 +3.2 +2.2 -2.9	2. 56 0. 81 10. 24 4. 84 8. 41	64. 2 63. 7 66. 9	-2.0 -1.7 -2.2 +1.0 +2.4	4. 00 2. 89 4. 84 1. 00 5. 76	11. 2 12. 8 14. 1	+2.2 0 +1.6 +2.9 +1.6	2.56 8.41	-1.53 -7.04 +2.20	+5.12 +6.38	-3.52 +2.90
1911 1912 1913 1914 1915	3. 8 6. 0 8. 1	-3.2 -3.0 -0.8 +1.3 +2.2	10. 24 9. 00 0. 64 1. 69 4. 84	64. 8 69. 6 67. 5	+7.5 -1.1 +3.7 +1.6 -5.5	56. 25 1. 21 13. 69 2. 56 30. 25	14. 2 9. 0 9. 0	+3.0 2.2 2.2	9.00 4.84 4.84	+3.30 -2.96 +2.08	+23. 04 -9. 00 +1. 76 -2. 86 +12. 76	-3.30 -8.14 -3.52
1916 1917		+1.4 -1.5	1.96 2.25		-4.4 -3.2	19. 36 10. 24		-4.9 +2.8	24. 01 7. 84	-6. 16 +4. 80	-6.86 -4.20	
Sums. Means		-0. 7 -0. 02			1. S 0. 07	237. 78				-9 4. 31		-151.31

$$Table \ 2-computations.$$

$$[p] = -0.7; \quad [t] = -1.8; \quad [y] = +1.6; \quad [py] = 82.87; \quad [ty] = -151.31; \quad [pt] = -94.31; \quad n = 27;$$

$$\sigma_p = \sqrt{\frac{[p^2] - [p]^2}{n}} = \sqrt{\frac{122.73 - (-0.7)^2}{27}} = \sqrt{\frac{122.73}{27}} = \sqrt{\frac{237.66}{27}} = \sqrt{8.80} = 3.0$$

$$\sigma_t = \sqrt{\frac{237.78 - (-1.8)^3}{27}} = \sqrt{\frac{251.04 - (\frac{11.6}{27})^3}{27}} = \sqrt{\frac{250.95}{27}} = \sqrt{9.29} = 3.0$$

$$\sigma_y = \sqrt{\frac{[py] - [p] [y]}{n}} = \frac{82.87 - (-0.7) (1.6)}{27} = \frac{82.87 + .04}{170.10} = \frac{82.91}{170.10} = +0.487$$

$$\sigma_t = \frac{-151.31 - (-1.8) (1.6)}{27} = -0.622.$$

$$\sigma_t = \frac{-94.31 - (-0.7) (-1.8)}{27} = -0.555$$

$$Er_{py} = \pm 0.6745 \frac{1 - r_{py}^2}{\sqrt{n}} = \pm \frac{0.6745}{5.2} \times 0.6131 = \pm 0.080$$

 $Er_{pt} = \pm \frac{0.6745}{5.2} \times 0.6920 = \pm 0.090$

NOMENCLATURE OF THE UNIT OF ABSOLUTE PRESSURE.

By CHARLES F. MARVIN.

[Weather Bureau, Washington, Mar. 30, 1918.]

While scientists are striving to secure international uniformity of units, nevertheless, right in our midst we find growing up a diversity of practice which all must deplore, regarding the nomenclature of pressure in absolute units, and which if not soon remedied will result in much future confusion. Pressure can be conceived only with reference to some area over which it acts, and pressure multiplied by area is a force. Since the dyne is the standard unit of force, a pressure of 1 dyne per square centimeter seems to constitute the logical unit of pressure. Indeed, since the concept pressure is inseparable from some area, science might be willing to grant that expressions like "a pressure of 1 dyne" has the same meaning as "a pressure of 1 dyne per square centimeter", etc., unless some other area is named. In other words, a unit of pressure is a unit of force on a unit area, and no particular name for this unit is really required. The practice and usage through the course of the last decade or more relative to the introduction of the names "barad", "barye", "bar", etc., as names for absolute units of pressure is briefly indicated in the following statement from notes furnished by C. F. Talman and Cleveland Abbe, jr., United States Weather Bureau.

A committee on uniformity of units of the British Association for the Advancement of Science recommended (see report of the Association for 1888, p. 28):

The unit of pressure on the C. G. S. system of units, i. e., the pressure of 1 dyne per square centimeter, to be called 1 barad.

At the International Physics Congress at Paris, 1900, M. Guillaume proposed that the name barye be applied to the megadyne per square centimeter. His proposal was referred to a committee on units, of which he was a member. This committee unanimously recommended that the name barye be applied to the dyne, instead of the megadyne, per square centimeter. The report of the committee was presented at the final session of the congress, but, so far as appears from the procès-verbaux, no action was taken on it.

T. W. Richards 1 and A. E. Kennelly 2 employed the term "bar" to signify a pressure of 1 dyne per square centimeter, having either selected the term independently or taken it as an abbreviated variation of the terms "barad" (British Association committee) or "barye" (Physicists' Congress, 1900). Other instances of restricted use of the terms barad and megabar are found; e. g., Tables of Physical and Chemical Constants, by Kaye and Laby (1911), page 5 and page 27, and Smith-sonian Physical Tables, fifth and sixth editions, pages 309 and 346, respectively.

Meteorologists have long had occasion to express atmospheric pressures in absolute terms,3 but it remained for Bjerknes to recognize the peculiar convenience in hydrodynamics and atmospherics of the megadyne per square centimeter as a unit of pressure, and, through his pupil J. W. Sandström, to introduce it—without assigning it a special name at the time—in his epochal investigations into the hydrodynamics of the sea and the atmos-

¹ Richards, T. W., in Carnegie Institution of Washington, Publication No. 7. Washington, 1903. pp. 42–43.

² Kennelly, A. E., in Proc. Am. instit. elect. eng., 1909, 28: 706.

³ e. g., Abbe, C. Preparatory studies for deductive methods in storm and weather predictions. Washington, 1890. p. 62. (Ann. rpt., C. S. O., 1889, App. 15.)

⁴ Sandström, J. W., & Helland-Hansen, B. Ueber die Berechnung von Meeresströmungen. Bergen, 1903. 8°. pp. 2, 14–15. (Report on Norwegian fishery and marine investigations, v. 2, 1902, No. 4.)